

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**

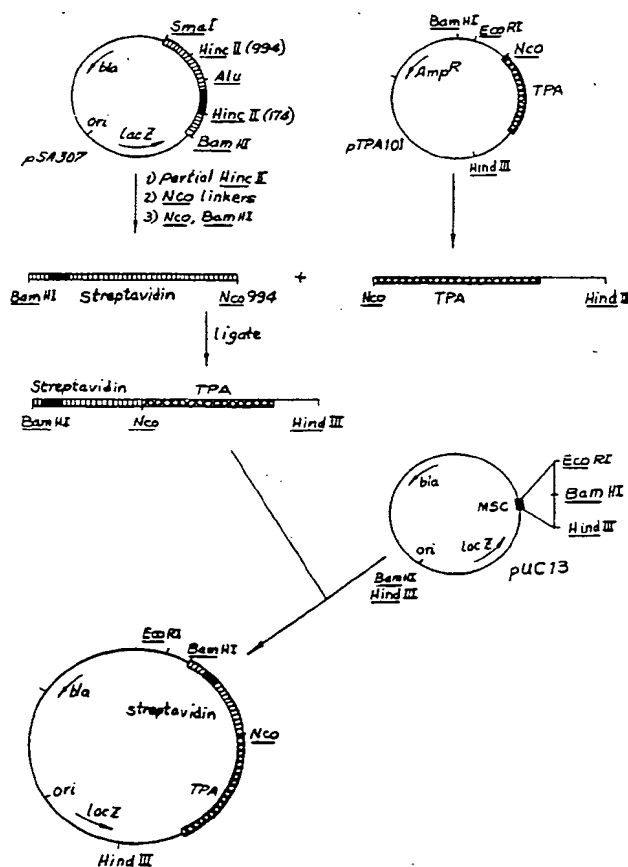


## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>4</sup> :</b> <b>C07H 21/04, C12N 1/00</b> <b>C07K 15/00, C12P 21/00, 21/02</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 86/02077</b> <b>(43) International Publication Date:</b> 10 April 1986 (10.04.86)
<b>(21) International Application Number:</b> PCT/US85/01901 <b>(22) International Filing Date:</b> 1 October 1985 (01.10.85) <b>(31) Priority Application Number:</b> 656,873 <b>(32) Priority Date:</b> 2 October 1984 (02.10.84) <b>(33) Priority Country:</b> US  <b>(71)(72) Applicants and Inventors:</b> MEADE, Harry, M. [US/US]; 62 Grasmere Street, Newton, MA 02158 (US). GARWIN, Jeffrey, L. [US/US]; 76 Fletcher Road, Bedford, MA 01730 (US). BIOGEN N.V. [NL/NL]; Pietermaai 15, Willemstad, Curaçao (AN).  <b>(74) Agent:</b> HALEY, James, F., Jr.; Fish & Neave, 875 Third Avenue, New York, NY 10022 (US).		<b>(81) Designated States:</b> AT (European patent), BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).  <b>Published</b> <i>With international search report.</i>

**(54) Title:** PRODUCTION OF STREPTAVIDIN-LIKE POLYPEPTIDES**(57) Abstract**

DNA sequences, hybrid DNA sequences, recombinant DNA molecules and processes for producing streptavidin-like polypeptides and for producing fused proteins consisting of a streptavidin-like polypeptide joined end to end with another protein, polypeptide, peptide or amino acid. The DNA sequences, hybrid DNA sequences and recombinant DNA molecules of this invention are characterized in that they include DNA fragments that code for streptavidin-like polypeptides. These DNA sequences, hybrid DNA sequences and recombinant DNA molecules and the hosts transformed with them may be employed in the processes of this invention to produce streptavidin-like polypeptides and fused proteins.



***FOR THE PURPOSES OF INFORMATION ONLY***

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

<b>AT</b> Austria	<b>GA</b> Gabon	<b>MR</b> Mauritania
<b>AU</b> Australia	<b>GB</b> United Kingdom	<b>MW</b> Malawi
<b>BB</b> Barbados	<b>HU</b> Hungary	<b>NL</b> Netherlands
<b>BE</b> Belgium	<b>IT</b> Italy	<b>NO</b> Norway
<b>BG</b> Bulgaria	<b>JP</b> Japan	<b>RO</b> Romania
<b>BR</b> Brazil	<b>KP</b> Democratic People's Republic of Korea	<b>SD</b> Sudan
<b>CF</b> Central African Republic	<b>KR</b> Republic of Korea	<b>SE</b> Sweden
<b>CG</b> Congo	<b>LI</b> Liechtenstein	<b>SN</b> Senegal
<b>CH</b> Switzerland	<b>LK</b> Sri Lanka	<b>SU</b> Soviet Union
<b>CM</b> Cameroon	<b>LU</b> Luxembourg	<b>TD</b> Chad
<b>DE</b> Germany, Federal Republic of	<b>MC</b> Monaco	<b>TG</b> Togo
<b>DK</b> Denmark	<b>MG</b> Madagascar	<b>US</b> United States of America
<b>FI</b> Finland	<b>ML</b> Mali	
<b>FR</b> France		

-1-

PRODUCTION OF  
STREPTAVIDIN-LIKE POLYPEPTIDES

TECHNICAL FIELD OF INVENTION

5           This invention relates to DNA sequences, recombinant DNA molecules and processes for producing streptavidin-like polypeptides and fragments thereof. More particularly, the invention relates to DNA sequences encoding streptavidin-like polypeptides and fragments thereof and recombinant DNA molecules containing those sequences for use in synthesizing those polypeptides and fragments in appropriate hosts, and in one embodiment secreting them and other products fused to them through the  
10           membrane of the host cell.

          In another embodiment of this invention these DNA sequences and recombinant DNA molecules may be used to produce fusion proteins with desired proteins, polypeptides, peptides, and amino acids  
20           by linking them to DNA sequences that code for those products and expressing the resulting hybrid gene in an appropriate host. Depending on the particular construction and streptavidin DNA fragment employed, these fused proteins may be secreted through the  
25           membrane of the host in which they are made. Again, depending on the particular construction and streptavidin fragment employed they may also be more easily purified because of the exceptionally strong

-2-

binding affinity of streptavidin for biotin and its derivatives.

#### BACKGROUND OF THE INVENTION

Streptavidin is an antibiotic produced by  
5 the bacteria Streptomyces avidinii and other  
Streptomyces species (E. O. Stapley et al.,  
Antimicrobial Agents and Chemotherapy 1963, 20-27  
(J. C. Sylvester ed. 1964)). It occurs naturally as  
a tetramer with a molecular weight of about 60,000  
10 daltons. Streptavidin is characterized by its strong  
affinity for biotin and biotin derivatives and  
analogues. Each of the four identical subunits has  
a single biotin binding site (K. Hofmann et al.,  
Proc. Natl. Acad. Sci. USA 77, 4666-68 (1980);  
15 L. Chalet and F. J. Wolf, Archives of Biochemistry  
and Biophysics 106, 1-5 (1964)).

Because of its strong affinity for biotin,  
streptavidin has found widespread application, both  
commercially, and for applied and basic biomedical  
20 research, to study biotin-requiring enzymes and, when  
used in conjunction with biotinylated substances, to  
study the interactions between these substances and  
other products (K. Hofmann, supra; see also E. A.  
Bayer and M. Wilchek, Methods of Biochemical Analysis  
25 26, 1-45 (1980); F.M.Finn et al., J.Biol Chem. 255,  
5742-46 (1980)). Today, streptavidin is produced  
commercially by isolating it from the cell medium of  
Streptomyces avidinii. Purification of streptavidin  
from Streptomyces avidinii has resulted in very low  
30 yields of only 3-4 mg per liter of cell culture.

Recent advances in molecular biology have  
made it possible to produce large amounts of heterol-  
ogous proteins in bacterial hosts. These include,  
for example, leukocyte interferon (S. Nagata et al.,  
35 "Synthesis In E. coli Of A Polypeptide With Human  
Leukocyte Interferon Activity", Nature 284, 316-20

-3-

(1980)), antigens of human hepatitis B virus (C. J. Burrell et al., "Expression In Escherichia coli Of Hepatitis B Virus DNA Sequences Cloned In Plasmid pBR322", Nature 279, 43-47 (1979) and M. Pasek et al., "Hepatitis B Virus Genes And Their Expression In E. coli", Nature 282, 575-79 (1979)), SV40 t antigen (T. M. Roberts et al., "Synthesis Of Simian Virus 40 t Antigen In Escherichia coli", Proc. Natl. Acad. Sci. USA 76, 5596-5600 (1979)), and FMD viral antigens (H. Kupper et al., "Cloning Of cDNA Of Major Antigen Of Foot And Mouth Disease Virus And Expression In E. coli", Nature 289, 555-59 (1982)).

In general, these processes rely on the construction of recombinant DNA molecules characterized by a DNA sequence coding for the desired protein, polypeptide, peptide or amino acid operatively linked to an expression control sequence. Appropriate hosts are then transformed with these molecules to permit production of the desired product by fermentation.

For DNA sequences, other than those prepared via chemical synthesis, the construction of such recombinant DNA molecules often comprises the steps of producing a single-stranded DNA copy ("cDNA") of a messenger RNA ("mRNA") template for the desired product; converting the cDNA to double-stranded DNA and operatively linking the DNA to an appropriate expression control sequence in an appropriate cloning vehicle. The recombinant DNA molecule is then employed to transform an appropriate host. Such transformation may permit the host to produce the desired product when it is fermented under appropriate conditions.

A further improvement of the above technology has made it possible to excrete the selected protein, polypeptide, peptide or amino acid through the membrane of the host cell as it is produced by:

forming a hybrid gene consisting of a DNA sequence from an extracellular or periplasmic

carrier protein that<sup>-4-</sup> is excreted by the host,  
and a heterologous DNA fragment which codes for  
the selected protein, polypeptide or amino acid;  
transforming the host with that hybrid  
5 gene operatively linked to an expression control  
sequence; and

culturing the transformed host to synthesize  
and to secrete the selected protein, polypeptide,  
peptide or amino acid.

10 Such techniques are disclosed, for example,  
by L. Villa-Komaroff et al., "A Bacterial Clone Syn-  
thesizing Pro-Insulin," Proc. Natl. Acad. Sci. USA 75,  
3727-31 (1978), and United States patent 4,411,994.  
However, any protein, polypeptide, peptide or amino  
15 acid made by this method, although separated from  
intracellular proteins and cell debris by secretion,  
must still be recovered from the cell medium or peri-  
plasmic space. This recovery generally involves a  
purification scheme that is less effective and less  
20 simple than desired. It also generally results in  
product losses.

#### SUMMARY OF THE INVENTION

The present invention solves the problems  
referred to above. It provides at least one DNA  
25 sequence characterized in that at least a portion  
thereof codes for a streptavidin-like polypeptide.  
The DNA sequences of this invention are selected  
from the group consisting of (a) SA304, SA307, SA324;  
(b) DNA sequences encoding polypeptides which  
30 hybridize to any of the foregoing DNA sequences and  
which code on expression for a streptavidin-like  
polypeptide; and (c) DNA sequences which code on  
expression for a polypeptide coded for on expression  
of any of the foregoing DNA sequences. These DNA  
35 sequences permit hosts transformed with them to pro-  
duce streptavidin-like polypeptides. The present

invention also provides<sup>-5-</sup> at least one hybrid DNA sequence characterized in that at least a portion thereof codes for a fused protein consisting of a streptavidin-like polypeptide joined end to end with  
5 tissue plasminogen activator ("TPA"). The DNA sequences of this aspect of the invention are selected from the group consisting of (a) SAT9724 SAT7021; (b) DNA sequences encoding polypeptides which hybridize to any of the foregoing DNA sequences and which  
10 code on expression for the fused protein; and (c) DNA sequences which code on expression for a polypeptide coded for on expression of any of the foregoing DNA sequences.

It is accordingly possible to use the DNA  
15 sequences, recombinant DNA molecules, hosts and processes of this invention to avoid the prior low yields which have beset other known methods of streptavidin production. Accordingly, they enable large amounts of highly pure streptavidin-like polypeptides  
20 and their derivatives to be made available for diverse uses in the pharmaceutical and other industries.

In another embodiment, this invention provides a method of producing, and secreting through  
25 the membrane of the host cell, a selected protein, polypeptide, peptide or amino acid by expressing a hybrid DNA sequence consisting of a DNA sequence coding for a sufficient portion of a prestreptavidin-like polypeptide to cause the resulting fused protein  
30 to be secreted through the membrane of the host cell and a DNA sequence coding for the selected protein, polypeptide, peptide or amino acid. Upon expression of this hybrid DNA sequence, the fused protein that is expressed from this hybrid DNA sequence is secreted  
35 through the membrane of the host cell transformed by the hybrid DNA sequence. It may then be used as a fusion protein or cleaved by a variety of chemical,

-6-

enzymatic and biological methods, to produce the desired protein, polypeptide, peptide or amino acid and a streptavidin-like polypeptide.

In still another embodiment, this invention provides a method for producing a fused protein that is purifiable by taking advantage of the binding affinity of streptavidin-like polypeptides for biotin and its derivatives or analogues. In this embodiment of the invention a hybrid DNA sequence consisting of a DNA sequence coding for a sufficient portion of a streptavidin-like polypeptide to cause the resulting fused protein to bind to biotin or its derivatives or analogues and a DNA sequence coding for the selected protein, polypeptide, peptide or amino acid is employed. Upon expression of this hybrid DNA sequence, the streptavidin moiety of the fused protein may be bound to biotin or one of its derivatives or analogues. Other secreted proteins or contaminants which do not bind to biotin can then be washed away and the fused protein eluted from the biotin. The selected protein, polypeptide, peptide or amino acid may then be used as a fusion protein or cleaved from the streptavidin-like polypeptide by conventional techniques, if necessary, and the streptavidin-like polypeptide and the selected protein, polypeptide, peptide or amino acid may then be collected separately.

In the most preferred embodiment of this invention, a fused protein is produced that is both secretable from the host cell in which it is made and purifiable by taking advantage of the binding affinity of streptavidin-like polypeptides for biotin and its derivatives or analogues. In this most preferred embodiment, a hybrid gene consisting of a DNA sequence coding for a sufficient portion of a pre-streptavidin-like polypeptide to cause the resulting fused protein to be secreted through the membrane of

-7-

the host cell and to cause the resulting fused protein to bind to biotin or its derivatives or analogues and a DNA sequence coding for the selected protein, polypeptide, peptide or amino acid is employed.

5

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic outline of one embodiment of making expression vectors pSA304 and pSA307, containing DNA sequences encoding the streptavidin-like polypeptides of this invention.

10

Figure 2 displays the nucleotide sequence of a portion of SA307 and the amino acid sequence of the regions of that portion coding for a streptavidin-like polypeptide. Line c of the amino acid sequence is the reading frame which is transcribed to produce a streptavidin-like polypeptide.

15

Figure 3 depicts an SDS-polyacrylamide gel used to determine the apparent molecular weight of the streptavidin-like polypeptides of this invention. Lane a is a sample of repurified commercial streptavidin; lane b is a sample of protein standards with bands, reading from top to bottom, of 43Kd, 25.7Kd, 18.4Kd, 14.3Kd, 12.3Kd, 6.2Kd and 3.0Kd; lane c is eco-avidin produced by expression of SA307 in E.coli; and lane d is a streptavidin-like polypeptide produced by expression of pSA3721 in S.lividans.

20

25

Figure 4 is a schematic outline of another embodiment of making an expression vector containing DNA sequences encoding the streptavidin-like polypeptides of this invention.

30

Figure 5 is a schematic outline of an embodiment of making an expression vector containing DNA sequences encoding a fused streptavidin-like-tissue plasminogen activator protein of this invention.

35

Figure 6 is a schematic outline of another embodiment of making an expression vector containing

DNA sequences encoding a fused streptavidin-like-tissue plasminogen activator protein of this invention.

Figure 7 is a schematic outline of yet another embodiment of making an expression vector containing DNA sequences encoding a fused streptavidin-like-tissue plasminogen activator protein.

#### DETAILED DESCRIPTION OF THE INVENTION

In accordance with this detailed description, the following definitions apply:

10        Protein -- A polypeptide containing a linear series of more than fifty amino acids, e.g., tissue plasminogen activator, pro-insulin, serum albumin, human growth hormone, parathyroid hormone, and interferon.

15        Polypeptide -- A linear series of amino acids connected one to the other by peptide bonds between the amino and carboxy groups of adjacent amino acids.

20        Precursor of a Protein or Polypeptide -- A polypeptide or protein as synthesized within a host cell with a signal sequence, e.g., pre-streptavidin, preproinsulin, preserum albumin, pregrowth hormone, preparathyroid hormone, and preinterferon. A mature polypeptide or protein is secreted through a host's cell membrane with the attendant loss or clipping (i.e., maturation) of the signal sequence of its precursor.

25        Nucleotide -- A monomeric unit of DNA or RNA consisting of a sugar moiety (pentose), a phosphate, and a nitrogenous heterocyclic base. The base is linked to the sugar moiety via the glycosidic carbon (1' carbon of the pentose) and that combination of base and sugar is a nucleoside. The base characterizes the nucleotide. The four DNA bases are adenine ("A"),

30

35

guanine ("G"), cytosine ("C") and thymine ("T"). The four RNA bases are A, G, C and uracil ("U").

5        DNA Sequence -- A linear series of nucleotides connected one to the other by phosphodiester bonds between the 3' and 5' carbons of adjacent pentoses.

10        Codon -- A DNA sequence of three nucleotides (a triplet) which encodes through messenger RNA ("mRNA") an amino acid, a translational start signal or a translational termination signal. For example, the nucleotide triplets TTA, TTG, CTT, CTC, CTA and CTG encode for the amino acid leucine ("Leu"), TAG, TAA and TGA are translational stop signals and

15        ATG is a translational start signal.

20        Plasmid -- A non-chromosomal double-stranded DNA sequence comprising an intact "replicon" such that the plasmid is replicated in a host cell. When the plasmid is placed within a unicellular host organism, the characteristics of that organism are changed or transformed as a result of the DNA of the plasmid. For example, a plasmid carrying the gene for tetracycline resistance (Tet<sup>R</sup>) transforms a host cell previously sensitive to

25        tetracycline into one which is resistant to it. A host cell transformed by a plasmid is called a "transformant".

30        Phage or Bacteriophage -- Bacterial virus, many of which include DNA sequences encapsidated in a protein envelope or coat ("capsid").

35        Cloning Vehicle -- A plasmid, phage DNA or other DNA sequence which is able to replicate in a host cell, characterized by one or a small number of endonuclease recognition sites at which its DNA sequence may be cut in a determinable fashion without attendant loss of an

-10-

essential biological function of the DNA, e.g., replication, production of coat proteins or loss of promoter or binding sites, and which contains a marker suitable for use in the identification of transformed cells, e.g., tetracycline resistance or ampicillin resistance. A cloning vehicle is also known as a vector.

Host -- An organism which, on transformation by a cloning vehicle, enables the cloning vehicle to replicate and to accomplish its other biological functions, e.g., the production of polypeptides or proteins through expression of the genes of a plasmid.

Cosmid -- A plasmid containing the cohesive end ("cos") site of bacteriophage  $\lambda$ . Cosmids may, because of the presence of the cos site, be packaged into  $\lambda$  coat protein and used to infect an appropriate host. Because of their capacity for large fragments of foreign DNA, cosmids are useful as cloning vehicles.

Expression -- The process undergone by a gene to produce a polypeptide or protein. It is a combination of transcription and translation.

Transcription -- The process of producing mRNA from a gene.

Translation -- The process of producing a protein or polypeptide from mRNA.

Promoter -- The region of the DNA of a gene at which RNA polymerase binds and initiates transcription. A promoter is located before the ribosome binding site of the gene.

Ribosome Binding Site -- The region of the DNA of a gene which codes for a site on mRNA which helps the mRNA bind to the ribosome, so that translation can begin. The ribosome binding site is located after the promoter and

-11-

before the translational start signal of the gene.

5        Gene -- A DNA sequence which encodes, as a template for mRNA, a sequence of amino acids characteristic of a specific protein, polypeptide or peptide.

10        Expression Control Sequence -- A DNA sequence that controls and regulates expression of genes of the cloning vehicle when operatively linked to those genes.

15        Signal DNA Sequence -- A DNA sequence within a gene for a polypeptide or protein which encodes, as a template for mRNA, a sequence of hydrophobic amino acids at the amino terminus of the polypeptide or protein, i.e., a "signal sequence" or "hydrophobic leader sequence" of the polypeptide or protein. A signal DNA sequence is located in a gene for a polypeptide or protein immediately before the DNA sequence coding for the mature protein as polypeptide and after the translational start signal (ATG) of the gene. A signal DNA sequence codes for the signal sequence of a polypeptide or protein, which (signal sequence) is characteristic of a precursor of the polypeptide or protein.

20        It is believed that only a portion of a signal sequence of a precursor of a protein or polypeptide is essential for the precursor of the protein or polypeptide to be transported through the cell membrane of a host and for the occurrence of proper clipping of the precursor's signal sequence to form the mature protein or polypeptide during secretion. Hence, the term "signal DNA sequence" means the DNA sequence which codes for the portion of the signal sequence essential to secretion and preferably to maturation of a precursor of a protein,

25       

30       

35

-12-

polypeptide or peptide, produced within a host cell.

5           Streptavidin-like Polypeptide -- A polypeptide which is substantially immunologically equivalent to natural streptavidin and is able to bind to biotin or biotin derivatives or analogues. This polypeptide may contain amino acids which are not part of natural streptavidin or may contain only a portion of natural streptavidin. The polypeptide may also not be identical to natural streptavidin because the host in which it is made may lack appropriate enzymes which may be required to transform the host-produced polypeptide to the structure of natural streptavidin.

10

15

#### THE HOST CELLS OF THIS INVENTION

Any of a large number of available and well-known host cells may be used in the host/expression vector combinations of this invention. The selection of a particular host is dependent upon a number of factors recognized by the art. These include, for example, compatibility with the chosen expression vector, toxicity to it of the proteins encoded for by the hybrid plasmid, ease of recovery of the desired protein, expression characteristics, bio-safety and costs. A balance of these factors must be struck with the understanding that not all hosts may be equally effective for the expression of a particular DNA sequence in the expression vectors and methods of this invention.

20

25

30

Within these general guidelines, useful microbial hosts may include strains of E.coli, Pseudomonas, Bacillus, Streptomyces, yeast and other fungi, insect, plant or animal (including human) cells in culture, or other hosts known in the art.

35

-13-

Most preferably, the host used in this invention is a gram positive bacterium such as Streptomyces. Gram positive bacteria do not possess an outer cell wall. Therefore, an excreted protein is transported across the cell membrane directly into the cell medium. This characteristic eliminates the need to isolate the protein from the periplasmic space, as often required when gram negative bacteria, such as E.coli, are used as host cells for protein secretion systems. The protein may be purified, free of all cellular contaminants, directly from the cell medium. The preferred gram positive host cell strain is Streptomyces lividans.

15                    THE EXPRESSION CONTROL  
                     SEQUENCES OF THIS INVENTION

In order to express the streptavidin DNA sequences or hybrid streptavidin-heterologous DNA sequences of this invention, those DNA sequences must be operatively linked to an expression control sequence. Methods of effecting this operative linking, either before or after the DNA sequence is inserted into the cloning vehicle, are well known.

Expression control sequences useful in this invention are also well known. They include the E.coli lac system, the E.coli trp system, the E.coli  $\beta$ -lac system, the TAC system, the TRC system, the major operator and promoter regions of bacteriophage  $\lambda$ , the control region of filamentaceous single-stranded DNA phages, the expression control sequences of Streptomyces or other gram positive bacteria and other sequences known to control the expression of genes of prokaryotic or eukaryotic cells and their viruses or combinations thereof.

More preferably, expression control sequences operative in gram positive bacteria are used. Most preferably, Streptomyces expression control

-14-

sequences are employed when Streptomyces host cells are used.

#### THE CLONING VEHICLES OF THIS INVENTION

5 In accordance with this invention, any cloning vehicle able to replicate in a host cell and having a restriction site into which DNA fragments can be inserted, may be used. The preferred cloning vehicles used in this invention are multicopy plasmids that are able to replicate in Streptomyces.

10 The specific location of the insertion site in the cloning vehicle into which the streptavidin DNA sequences or hybrid streptavidin-heterologous DNA sequences of this invention are inserted, is not critical for making the selected  
15 protein, polypeptide, peptide or amino acid. In this regard, the cleavage site can be located anywhere in the cloning vehicle that does not interfere with its replication in the host cell.

Restriction enzymes used to create the  
20 desired cleavage site are well known. They include, for example, AraI, PstI, SalI, EcoRI, BamHI, HindIII, HincII and Sau3a. Methods for cleaving the cloning vehicles used in the methods of this invention at the desired restriction site and inserting into that  
25 site a DNA sequence are also well-known.

#### EXPRESSION OF DNA SEQUENCES ACCORDING TO THE METHODS OF THIS INVENTION

The methods of this invention may be employed to express DNA sequences coding for a streptavidin-like  
30 polypeptide or hybrid DNA sequences consisting of a DNA sequence coding for a portion of a streptavidin-like polypeptide and a DNA sequence coding for a desired eukaryotic, prokaryotic or viral protein, polypeptide, peptide or amino acid in a variety of ways.

-15-

In the embodiment of this invention in which it is desired to produce streptavidin-like polypeptides alone, the DNA sequence coding for the desired polypeptide is operatively-linked to an expression control sequence in a cloning vehicle. In such a construction, there should be a translational start signal immediately preceding the DNA sequence coding for the streptavidin-like polypeptide, if that sequence does not itself begin with a start signal. There also should be no stop codon between the start signal and the end of the DNA sequence coding for the streptavidin-like polypeptide. The resulting recombinant DNA molecule is then used to transform an appropriate host and the transformed host cultured under conventional fermentation conditions to produce the desired streptavidin-like polypeptide.

In a preferred embodiment of this aspect of the invention, a DNA sequence encoding a portion of a carrier protein, including a sufficient portion of the carrier protein signal sequence to cause the fused carrier protein-streptavidin-like polypeptide to be secreted from the cell in which it is made, resides between the translational start signal and the DNA sequence encoding the streptavidin-like polypeptide. More preferably only the carrier protein signal DNA sequence, without any of the DNA sequences coding for the carrier protein itself, is present between the translational start signal and the DNA sequences coding for the streptavidin-like polypeptide. Most preferably, the signal DNA sequence present between the translational start signal and the DNA sequences encoding a streptavidin-like polypeptide is the streptavidin signal DNA sequence, a sufficient portion of which is present to cause the streptavidin-like polypeptide to be secreted from the host cell.

-16-

In this manner, the streptavidin-like polypeptide is produced and secreted from the transformed host. Preferably, the signal sequence is cleaved from the streptavidin-like polypeptide or the fused carrier protein-streptavidin-like polypeptide during its secretion from the cell. In the embodiment where a portion of a carrier protein is fused to the streptavidin-like polypeptide, after production and isolation of the fused protein, it is preferable to cleave the streptavidin-like polypeptide from the carrier protein using known methods discussed infra.

In the embodiment of this invention in which it is desired to produce a non-streptavidin-like protein, polypeptide, peptide or amino acid using the methods of this invention, the DNA sequence coding for that product is linked downstream from and in the same reading frame as at least a portion of the DNA sequence coding for a streptavidin-like polypeptide to form a hybrid DNA sequence. The resulting hybrid DNA sequence is then operatively linked to an expression control sequence in a cloning vehicle. Again, in such a construction there should be a translational start signal immediately preceding the hybrid DNA sequence, if the hybrid DNA sequence does not itself begin with a start signal. There also should be no stop codon between the start signal and the end of the hybrid DNA sequence.

In various embodiments of this aspect of the invention, some or all of the DNA sequences encoding the carrier protein/carrier protein signal sequence/streptavidin signal sequence combinations described supra, reside between the translational start signal and the hybrid DNA sequence. In this manner a fused protein consisting of the streptavidin-like polypeptide joined end to end with the non-streptavidin-like protein, polypeptide, peptide or amino acid is produced and secreted from the transformed host. In

another embodiment, the streptavidin signal DNA sequence is directly linked to the DNA sequence encoding the non-streptavidin-like protein, polypeptide, peptide or amino acid to allow secretion of the protein, polypeptide, peptide or amino acid from the host cell. In all the above embodiments, the signal sequence preferably is cleaved upon secretion from the host cell. In the embodiment in which a portion of a carrier protein is fused to the fused protein, it is preferable to cleave the fused protein from the carrier protein using known methods discussed infra after production and isolation of the fused protein.

The appropriate portion of the DNA sequence coding for a streptavidin-like polypeptide or pre-streptavidin-like polypeptide employed is determined by a number of factors. These include the expression characteristics of the DNA sequence encoding the desired product, the ease of secretion of the desired product, the ultimate use to which the desired product is to be put, and whether secretion using the streptavidin sequences, purification using streptavidin binding or both are desired.

For example, if the desired product is to be secreted using the streptavidin signal sequence and purified using streptavidin binding to biotin or biotin derivatives or analogues, a sufficient portion of the streptavidin signal sequence to allow secretion, and a sufficient portion of the mature streptavidin coding sequence to allow binding to biotin or one of its derivatives or analogues is required.

It should also be understood that if the non-streptavidin-like protein, polypeptide, peptide or amino acid itself is required, the streptavidin-derived sequences may be removed after production, secretion, purification or any combination of them

-18-

by a variety of means. For example, the fused protein may include a chemical or enzymatic cleavage site useful to separate the desired product from the streptavidin-like polypeptide. In another embodiment of this invention, such a cleavage site is built into the fused protein by constructing the hybrid DNA sequence so that it has one or more codons between the portion coding for the undesired part of the fused protein and the portion coding for the desired part. On expression, these codons produce the desired cleavage site.

Because streptavidin contains no methionine residues, a preferred means of cleaving the fused protein -- provided the desired protein, polypeptide, peptide or amino acid also contains no methionine residues -- is to construct the hybrid DNA sequence with an ATG codon (coding for methionine) between the DNA sequence encoding the streptavidin-like polypeptide and the DNA sequence encoding the desired product. The fused protein may then be cleaved at the lone methionine residue by treatment with cyanogen bromide. (See E. Gross, Methods In Enzymology 11, 238-55 (1967).)

Among the heterologous DNA sequences which are useful in this invention are those which code for animal and human hormones, such as any of the various IFN- $\alpha$ 's, particularly  $\alpha 2$ ,  $\alpha 5$ ,  $\alpha 7$ ,  $\alpha 8$ , IFN- $\beta$ , IFN- $\gamma$ , human insulin and growth hormone, bovine growth hormone, swine growth hormone and erythropoietin, human blood factors and tissue plasminogen activator, viral or bacterial antigens, such as the core or surface antigen of HBV or the antigens of FMDV and other useful polypeptides of prokaryotic, eukaryotic or viral origin. Preferably, the heterologous DNA fragment used also contains its gene's translational stop signal and most preferably a part of its 3' non-coding region.

PURIFICATION OF THE PRODUCTS OF THIS INVENTION

The more advantageous embodiments of this invention are those in which the products produced, either a streptavidin-like polypeptide or a fused protein having a sufficient portion of a streptavidin-like polypeptide to permit binding to biotin or one of its derivatives or analogues, are purified by binding to biotin or a biotin derivative or analogue. The unbound contaminants are then discarded and the desired product recovered from the biotin. Of course, biotin-binding can also be used to separate the streptavidin residue of a fused protein from the desired product after cleavage of the fused protein as described previously.

Preferably, the streptavidin-like polypeptides of this invention are purified using bioaffinity chromatography. Most preferably, the chromatography resin is formed by covalently binding imminobiotin, a biotin analogue, to agarose, or some other inert matrix (see Hofmann et al., supra; G. Heney and G. A. Orr, Anal. Biochem. 114, 92-96 (1981)).\*

Streptavidin binds to imminobiotin at basic pH and is eluted at acidic pH in the presence of high concentrations of urea. Most preferably, the streptavidin-like polypeptide is applied to an imminobiotin-agarose column in 5 mM sodium carbonate buffer,

---

\* Streptavidin binds so strongly to unmodified biotin-agarose resin that it can be eluted only under very harsh conditions (see P. Cuatrecasas and M. Wilchek, Biochem. Biophys. Res. Commun. 33, 235-39 (1968)). Therefore, biotin agarose is unsatisfactory for use in purifying streptavidin-like polypeptides. However, biotin-agarose or similar resin may be useful in purifying a fusion protein of a streptavidin-like polypeptide linked to a selected protein, polypeptide, peptide or amino acid, which, because of the presence of the selected protein, polypeptide, peptide or amino acid, has a lower affinity for biotin.

-20-

pH 10.5, 0.5 M NaCl, and eluted with 50 mM sodium acetate buffer, pH 3.8, 1 M urea.

In order that this invention be more fully understood, the following example, for illustrative purposes only, is included herein.

#### EXAMPLE

##### PREPARATION OF A STREPTAVIDIN DNA PROBE

We determined the amino acid sequence of the first 38 N-terminal amino acids of commercial streptavidin. (purchased from BRL). We then chemically synthesized a 14-base probe ("SA-1") corresponding to a portion of the amino terminus of streptavidin (trp-tyr-asn-gln-leu).<sup>\*</sup> This DNA probe was used to screen the cosmid library, prepared as described below, for DNA sequences encoding streptavidin-like polypeptides.

##### PREPARATION OF AN S. AVIDINII COSMID LIBRARY

We cultured S. avidinii cells in YME media and isolated the bacterial DNA in a standard manner (K. F. Chater et al., Current Topics In Micro. and Immunol. 96, 69-75 (1982)). We then partially digested the DNA with Sau3A and size fractionated it by centrifugation through a salt gradient. We collected fragments larger than 20kb for insertion into the cosmid pHC79.

The cosmid pHC79, sold commercially by Boehringer Mannheim, consists of a portion of the plasmid pBR322, including the gene coding for ampicillin resistance, and the "cos" region, coding for

\* The SA-1 probe was 16-fold degenerate and consisted of the sequence:

AACATATTAGTTGA  
G G CA

-21-

the complementary ends of bacteriophage lambda DNA (B. Hohn and J. Collins, Gene 11, 291-98 (1980)).

We linearized pHC79 with BamHI, which cleaves the plasmid at a single site and produces  
5 ends complementary to those of the Sau3A-digested DNA. We then ligated the S. avidinii fragments to the linearized pHC79. The resultant recombinant DNA molecules, comprising the S. avidinii cosmid library, were added to a lambda in vitro packaging extract  
10 (Amersham). We used the resulting lysate to transfect E.coli K-12 strain ED8767 by standard procedures (Maniatis et al., Molecular Cloning, A Laboratory Manual, 295-305 (Maniatis, Fritsch and Sambrook, ed. 1982)).

15                    SCREENING OF THE S. AVIDINII  
                    COSMID LIBRARY WITH THE SA-1 PROBE

To screen the above-described library for DNA sequences encoding streptavidin, we grew the transfected E.coli cells in LB medium supplemented  
20 with ampicillin (50 µg/ml) at 37°C. Because cosmid pHC79 includes the gene coding for ampicillin resistance, E.coli K-12 cells which contain pHC79 will grow in the presence of ampicillin, while K-12 cells which do not contain pHC79 will not be able to grow.  
25 Therefore, growth in ampicillin-containing medium permits selection of hosts containing pHC79.

We selected 2200 Amp<sup>R</sup> colonies and picked them into individual wells of 96-well microtiter plates. We grew the selected cultures as before at  
30 37°C and printed them out onto nitrocellulose filters in an array using a 96-pronged fork. We then hybridized the colonies with the SA-1 probe, prepared as described above, at 30°C in 6XSSC buffer, 0.1% SDS overnight, using the hybridization methods of  
35 Grunstein and Hogness, Proc. Natl. Acad. Sci. USA 72, 3961-65 (1975). One of the 2200 colonies

-22-

contained DNA sequences which hybridized to the SA-1 probe (colony 62/c).

#### ISOLATION OF THE STREPTAVIDIN-RELATED DNA

We isolated the plasmid DNA from colony  
5 62/c and reconfirmed its hybridization to the SA-1 probe in the same manner as above. The plasmid DNA (p62/c) contained an insert of S. avidinii DNA of approximately 40kb at the BamHI site of pHC79. Referring now to Figure 1, we digested this 40kb  
10 fragment with BamHI and isolated a 2kb fragment which also hybridized to SA-1. We designated this 2kb fragment SA304. We then linearized plasmid pUC8 (J. Viera and J. Messing, Gene 19, 259-68 (1982)) with  
15 the BamHI, which cleaves at one site, downstream from the lac promoter. We ligated the linearized pUC8 to SA304 in a conventional manner. This ligation produced a recombinant DNA molecule which we designated pSA304.

We found that E.coli cells transformed  
20 with pSA304 produced and secreted across the cell membrane streptavidin-like polypeptides. However, this production and secretion occurred only when SA304 was in an orientation such that it was operatively linked to the lac promoter present on pUC8.  
25 On the basis of these results we concluded that pSA304 contained the E.coli lac promoter operatively linked to and controlling expression of DNA sequences encoding a streptavidin-like polypeptide, including the signal DNA sequences.\*

30 We confirmed the presence on SA304 of DNA sequences coding for a streptavidin-like polypeptide

---

\* We confirmed the presence of the streptavidin signal DNA sequence by nucleotide sequencing of SA307,  
35 infra. We further confirmed the presence of the S.avidinii streptavidin expression control sequence by S1 mapping described infra.

-23-

by sequencing portions of pSA304. The nucleotide sequence near nucleotide number 623 (Alu) (see Figure 2), within the SA304 fragment, matched exactly that of the SA-1 probe. The nucleotide sequence also conformed to that predicted from the protein sequence of a 20 amino acid region of the streptavidin polypeptide downstream from the region corresponding to the SA-1 probe.

10                    DETERMINING THE NUCLEOTIDE SEQUENCE  
                    OF THE STREPTAVIDIN-LIKE POLYPEPTIDE

Referring again to Figure 1, in order to localize more precisely and sequence the DNA fragment encoding the streptavidin-like polypeptide, we treated pSA304 with BamHI. We reisolated fragment SA304 and ligated it to the plasmid pUC13 which had been linearized with BamHI.\* We next restricted the resultant plasmid (pSA306) with SmaI and religated the large fragment, resulting in the loss of a 900 base pair fragment of SA304 to form plasmid pSA307. We then treated pSA307 with BamHI and SmaI and isolated a 1.1Kb fragment ("SA307"). We sequenced this fragment using the method of Maxam and Gilbert, Methods In Enzymology 65, 499-560 (1980). The sequence (Figure 2) indicates the presence of a DNA fragment coding for a streptavidin-like polypeptide between the ATG translational start signal at nucleotide number 480 and the TAG stop signal at nucleotide number 1030. The sequence also indicates the presence of a signal DNA sequence coding for a 35 amino acid signal sequence. We confirmed the presence of the

---

\* Plasmid pUC13 is identical to pUC8 except that the nucleotides immediately surrounding the BamHI site are in a different orientation and contain an additional restriction site (J. Viera and J. Messing, supra).

-24-

streptavidin expression control sequence by S1 mapping, described infra.

EXPRESSION OF  
STREPTAVIDIN-LIKE POLYPEPTIDES IN E.COLI

5           We employed pSA304 and pSA307 to transform E.coli K-12 cells using standard procedures.

          We then prepared and tested cultures of these transformed organisms for the expression and secretion into the periplasmic space of streptavidin-  
10   like polypeptides by the immunological reactivity,\* biotin-binding activity and amino acid sequence of the N-terminal end of the E.coli-produced polypeptide.

          We cultured strain JM83 of E.coli K-12 cells transformed with pSA304 in LB medium containing  
15   40 µg/ml ampicillin at 37°C for approximately 18 hours. We isolated proteins from the periplasmic space using a cold osmotic shock technique (L. A. Heppel, Methods In Enzymology 126, 841-47 (1968)). Rabbit antiserum raised against the commercial strep-  
20   tavidin as described above recognized two major periplasmic polypeptides isolated from the osmotic shock fluid on western blots (see H. Towbin et al., supra). The larger of the two streptavidin-like polypeptides had an apparent molecular weight of  
25   about 17,500 daltons, determined by SDS-polyacrylamide gel electrophoresis, as compared with an apparent molecular weight of commercial streptavidin of 13,500 daltons (Figure 3). The smaller of the two polypeptides had an apparent molecular weight of

30

\*       We raised rabbit anti-streptavidin antibodies against commercially-available streptavidin which we repurified by P-150 gel filtration chromatography. The antiserum recognized the commercial streptavidin and streptavidin purified from S. avidinii cell  
35   medium, as judged by immunoblots of protein electrophoretically transferred from SDS-polyacrylamide gels to nitrocellulose sheets (see H. Towbin et al., Proc. Natl. Acad. Sci. USA 76, 4350-54 (1979)).

-25-

15,000 daltons. The two streptavidin-like polypeptides co-eluted from iminobiotin-agarose chromatography columns. We next subjected the E.coli-produced streptavidin-like polypeptides ("eco-avidin major" and "eco-avidin minor") to P-150 gel filtration chromatography under non-denaturing conditions in 10mM tris-HCl buffer, pH 8.0, 0.9% NaCl. The two eco-avidins migrated with an apparent molecular weight of approximately 88,000 daltons. This molecular weight is most consistent with a native tetrameric structure, similar to the naturally-occurring streptavidin.

We subjected the eco-avidins to biotin-agarose and Sepharose CL-6B (unmodified agarose) column chromatography in the presence of 10 mM tris-HCl buffer, pH 8, 0.9% NaCl. The eco-avidins bound to the biotin-agarose, but did not bind to unmodified agarose. The eco-avidins also bound to columns of iminobiotin and displayed an elution pattern similar to that displayed by native streptavidin (see K. Hofmann et al., Proc. Natl. Acad. Sci. USA 77, 4666-68 (1980)).

We further confirmed the identity between naturally-occurring streptavidin and both forms of eco-avidin by sequencing the amino terminus of the eco-avidin polypeptides using an Applied Biosystems gas-phase amino acid sequenator. We separated the major and minor forms of the eco-avidin using SDS-polyacrylamide gel electrophoresis, electroeluted the polypeptides and separately determined the amino acid sequence of the amino terminus of each. The eco-avidins contained 13 amino acids at their amino terminus not present in naturally-occurring, mature streptavidin, followed by 12 amino acids identical

-26-

to those of naturally-occurring streptavidin.\* The protein sequence agreed precisely with that predicted from the nucleotide sequence of the corresponding region of SA307 and confirmed the cleavage in E.coli of a 22 amino acid fragment of the streptavidin signal sequence. Thus a 13 amino acid fragment of the streptavidin signal sequence remains in both forms of mature eco-avidin.

We assayed the eco-avidin by its ability to bind radioactively-labelled biocin (see S. G. Korenman and B. W. O'Malley, Meth. Enzymol. 18A, 427-30 (1970)). The yield of purified eco-avidin was about 30 mg/liter of culture medium, a 7.5-10 fold increase over the yield of naturally-occurring streptavidin prepared from S. avidinii. E.coli cells transformed with pSA307 and assayed as above produced similar levels of streptavidin.

We confirmed in the same manner as with pSA304, supra, that the expression control sequence controlling expression of the streptavidin-like DNA fragments in pSA307-transformed E.coli cells was the E.coli lac promoter.

#### EXPRESSION OF STREPTAVIDIN- LIKE POLYPEPTIDES IN S. LIVIDANS

We have also cloned DNA sequences encoding a streptavidin-like polypeptide under the control of the streptavidin expression control sequence into an expression vector which is able to replicate in the gram positive bacteria Streptomyces lividans. Since

30

\* Since both eco-avidin major and eco-avidin minor have the same amino terminus but different apparent molecular weights, they must differ at the carboxyl terminus. Upon hydroxylamine cleavage, eco-avidin major is converted to a form with an apparent molecular weight similar to that of eco-avidin minor. This is consistent with the inference that the two forms differ only at the carboxyl terminus.

35

-27-

pSA304 did not replicate in S.lividans, we made the cointegrant plasmid consisting of pSA304 and pIJ702. In this plasmid the streptavidin DNA sequences are operatively linked to both a lac promoter, located  
5 directly upstream from SA304, the inserted DNA sequence, and the streptavidin expression control sequence, present on SA304.

Referring now to Figure 4, we linearized pSA304 with PstI and treated the linearized plasmid  
10 with calf intestinal phosphatase to prevent recircularization. We next linearized pIJ702 (E. Katz et al., J. Gen. Micro. 129, 2703-14 (1983)) with PstI and ligated it to the linearized pSA304. The plasmid pIJ702 contains the Streptomyces marker gene  
15 for thiostrepton resistance as well as a replicon which allows growth in Streptomyces host cells.

We transformed E.coli JM83 cells with the ligation mixture and incubated as previously described in the presence of ampicillin. We isolated one Amp<sup>R</sup>  
20 colony containing the cointegrant plasmid pSA3721.

We isolated pSA3721 and used it to transform Streptomyces lividans. We selected for thiostrepton resistance and isolated eight colonies. We cultured each colony for up to eight days in R2YE medium containing 2 µg/ml thiostrepton in a baffled shaker  
25 flask at 30°C. We harvested the cells and culture medium at various times and, using the immunological and biotin binding assays described above, we tested for the production, and secretion into the cell  
30 medium, of a streptavidin-like polypeptide. We found that the transformed S. lividans produced, and secreted into the cell medium, a polypeptide which behaved in these assays identically to natural streptavidin. The level of production was approximately  
35 250 mg. protein per liter of culture medium, an increase over S. avidinii production of about 60-80

-28-

fold. We confirmed that the streptavidin-like DNA sequences are under the control of the streptavidin expression control sequence present in the SA304 fragment by performing S1 mapping experiments (A. J. Berk and P. A. Sharp, Cell 12, 721-32 (1977)) which show that the RNA transcript begins at the same site in S. avidinii and pSA3721.

In order to confirm that the streptavidin-like polypeptide produced and secreted by S. lividans was the result of expression of DNA fragments contained on the pSA3721 recombinant DNA molecule, we transformed S. lividans with pIJ702 alone under identical conditions and selected for thiostrepton resistance as above. None of the colonies isolated in this manner produced a protein with streptavidin-like activity as determined by any of the above-described assays.

#### PRODUCTION OF A FUSED PROTEIN OR POLYPEPTIDE

We next construct a hybrid DNA sequence by inserting into plasmid pSA307 DNA sequences encoding a selected protein, polypeptide, peptide or amino acid at a restriction site located at the end but before the translational stop signal of, and in the same reading frame as, the DNA sequences encoding the streptavidin-like polypeptide.

To effect this construction, for example, we use the DNA sequence encoding  $\alpha$ -antitrypsin (" $\alpha$ AT"). We treat pULB1523 (A. Bollen et al., Gene 2, 255-64 (1983)) with BamHI and Pst and isolate the small DNA fragment which encodes all but the first two amino acids of  $\alpha$ AT. We add a synthetic linker to the 5' end of this fragment to reconstruct the codons coding for the first two amino acids and to add an Nco restriction site at the 5' end of the fragment. We next linearize pSA304 by partial cleavage with HincII. We add Nco linkers to the

-29-

linearized pSA304 and cleave with Nco to create cohesive ends. We next treat the linearized pSA304 with Pst. We select the fragment which has been cleaved by HincII and has an Nco linker only at nucleotide  
5 994 of SA304. We ligate the DNA sequence encoding  $\alpha$ AT to this DNA fragment to create a recombinant DNA molecule having the  $\alpha$ AT inserted adjacent to and in the same reading frame as the DNA sequences encoding the streptavidin-like polypeptide.

10 We transform E.coli K12 cells with the resultant recombinant DNA molecule and culture the transformed host cells to produce and secrete through the host cell membrane, a fused polypeptide consisting of the streptavidin-like polypeptide joined  
15 end-to-end with the  $\alpha$ AT.

Other DNA sequences are also available and may be used in the processes of this invention to produce a fused protein consisting of a streptavidin-like polypeptide joined to a selected protein, polypeptide, peptide or amino acid.  
20

To construct a cloning vehicle which is able to replicate in Streptomyces and produce a fused streptavidin-like- $\alpha$ AT protein, we linearize the recombinant DNA molecule containing the hybrid DNA  
25 sequence with PstI and treat the linearized recombinant DNA molecule with calf intestinal phosphatase to prevent recircularization. We next ligate PstI-linearized pIJ702 to the linearized recombinant DNA molecule to create a cointegrant plasmid containing  
30 the hybrid streptavidin- $\alpha$ AT DNA sequence. We transform S. lividans cells with the cointegrant plasmid and culture the host cells to produce and secrete through the host cell membrane and into the cell medium a fused protein consisting of a streptavidin-like polypeptide joined end to end with  $\alpha$ AT.  
35

-30-

In another embodiment of this invention, we constructed a hybrid DNA sequence coding for a fused protein consisting of a streptavidin-like polypeptide fused end to end with tissue plasminogen activator ("TPA").

Referring now to Figure 5, we first constructed a plasmid containing DNA sequences encoding a streptavidin-like polypeptide. We treated pSA307 with HincII which cuts at nucleotide numbers 174 and 994 within SA307 (Figure 2). We next added Nco linkers to both ends of this fragment and cut with Nco to create cohesive ends. We designated this fragment SA324.

We then linearized plasmid pKK233-2 (J. Brosius, personal communication) with Nco. We ligated the linearized pKK233-2 to SA324 to produce plasmid pSA324. We transformed E.coli K12 cells with pSA324 and cultured the transformed host cells to produce and secrete through the host cell membrane, a streptavidin-like polypeptide. Production of the streptavidin-like polypeptide was induced by IPTG. From this we concluded that pSA324 contains DNA sequences encoding a streptavidin-like polypeptide, including its signal DNA sequence, operatively linked to the TRC expression control sequence present on pKK233-2.

We next constructed a plasmid containing sequences encoding TPA. Plasmid pTPA20 contains a DNA sequence encoding mature TPA inserted between a BglII restriction site at the 3' end and a HindIII restriction at the 5' end of the TPA gene. We treated pTPA20 with BglII and HindIII and isolated the TPA fragment. We ligated an Nco-BglII linker to the 5' end of the TPA gene. We next treated pKK233-2 with Nco and HindIII. We isolated the large fragment and ligated it to the TPA fragment ("pTPA101"). We next cut pTPA101 with BamHI and inserted a fragment

-31-

containing the lacI (TRC repressor) gene.\* We designated this plasmid pTPA102.

We constructed the hybrid streptavidin-TPA DNA sequence as follows. We treated pSA324 with  
5 Nco to isolate SA324. We next linearized pTPA102 with Nco and treated with calf intestinal phosphatase to prevent recircularization. We ligated the linearized pTPA102 to SA324. We designated this plasmid pSAT9724 and the hybrid DNA sequence SAT9724 present  
10 therein.

We transformed E.coli HB101 with pSAT9724 and cultured the transformed host. We tested for production of a fused streptavidin-like-TPA protein using the immunological assay described above  
15 (H. Towbin et al., Proc. Natl. Acad. Sci. USA 76, 4350-54 (1979)). We found that the transformed E.coli produced a polypeptide which contained both streptavidin and TPA immunologically reactive material. The protein had an apparent molecular  
20 weight of approximately 80,000 daltons. This is consistent with the molecular weight of a fused streptavidin-TPA protein. Production of the fused protein was induced by IPTG, indicating that the hybrid DNA sequence was under the control of the  
25 TRC promoter present on pKK233-2.

We next constructed a plasmid which is able to replicate in S.lividans. We cut pSAT9724 with BamHI and ligated it to pIJ702 which had been linearized with BglII to form cointigant plasmids  
30 pSAT9786 and pSAT9790. These plasmids are used to

---

\* The lacI gene codes for the repressor of the TRC promoter. In the presence of the lacI repressor the TRC promoter will only initiate transcription  
35 upon addition of an inducer such as IPTG, thus, enabling us to turn transcription on or off by controlling the level of IPTG added to the host cell.

-32-

transform S.lividans which is then cultured to produce and secrete into the cell medium a fused protein consisting of a streptavidin-like polypeptide joined end to end with TPA.

5 In yet another embodiment of this invention, we constructed a recombinant DNA molecule containing a hybrid DNA sequence which produced a fused protein, consisting of a streptavidin-like polypeptide fused end to end with TPA, only in Streptomyces.

10 Referring now to Figure 6, we treated pSA324 with Nco and reisolated SA324 containing the DNA sequence encoding a streptavidin-like polypeptide. We linearized pSAT7001, which contains the DNA sequence encoding TPA, with Nco, which cuts in  
15 front of the TPA gene, and ligated it to SA324. We isolated plasmid pSAT7020. We confirmed by nucleotide sequencing that the streptavidin gene was in the proper orientation in front of the TPA gene to form a hybrid DNA sequence encoding a fused protein  
20 consisting of a streptavidin-like polypeptide joined end to end with TPA.

In order to reconstruct the first 174 nucleotides of SA304 removed upon construction of SA324, we treated pSA307 with BamHI and with BssH2,  
25 which cleaves at nucleotide 422 of SA304 (see Figure 2). We isolated the small fragment. We next treated pSAT7020 with BamHI and BssH2 and ligated the large fragment to the small pSA304 fragment. We designated the resultant plasmid pSAT7021 and the  
30 hybrid DNA sequence SAT7021 present therein.

We next linearized pSAT7021 with BamHI and treated the linearized plasmid with calf intestinal phosphatase to prevent recircularization. We ligated BglII-linearized pIJ702 to the linearized pSAT7021  
35 to create a cointegrant plasmid pSAT7026 containing the hybrid streptavidin-TPA DNA sequence. We transformed S.lividans cells with pSAT7026. We isolated

-33-

from the cell medium a fused protein, containing immunological activity, as determined by the assay described above, of both streptavidin and TPA.

Although we constructed pSAT7021 as described, a similar plasmid could be more easily constructed as follows. Referring now to Figure 7, we cleave pSA307 with HincII under conditions in which there is only partial HincII cleavage. We add Nco linkers and cleave with Nco and BamHI. We isolate the 994 nucleotide fragment containing most of the streptavidin gene. We cleave pTPA101 with Nco and HindIII and isolate the DNA fragment encoding TPA. We ligate the two fragments together to produce a hybrid DNA sequence consisting of the streptavidin gene joined end to end with, and in the same reading frame as, the TPA gene. We next cleave pUC13 with BamHI and HindIII and ligate the large fragment to the hybrid DNA sequence. This produced a plasmid containing the hybrid streptavidin-TPA gene under the control of the streptavidin expression control sequence. This plasmid can then be used to form a cointegrant plasmid able to replicate in Streptomyces.

We next isolate secreted proteins from the cell medium using standard techniques [e.g., ammonium sulfate precipitation] and subject the isolate to imminobiotin-agarose column chromatography as described above to separate the fused protein from any contaminant proteins. We elute the fused protein from the imminobiotin-agarose with high pH-urea buffer as described supra and cleave the streptavidin-like polypeptide from the selected protein, polypeptide, peptide or amino acid using a known technique in a manner that does not cleave within either the streptavidin or the selected protein or polypeptide. For example, if the desired protein does not include a methionine, cyanogen bromide may be employed as a cleavage agent. Other methods of protein cleavage

have been described supra and may be used in the processes of this invention.

We then separate the streptavidin-like polypeptide from the selected protein, polypeptide, peptide or amino acid by various means, preferably using imminobiotin-agarose column chromatography. The selected protein, polypeptide, peptide or amino acid passes through the column without binding to the imminobiotin and the streptavidin-like polypeptide is eluted from the column as above. If the streptavidin-like polypeptide and the selected protein, polypeptide, peptide or amino acid are of sufficiently different molecular weights, they may also be separated by one of several sizing techniques known in the art such as native polyacrylamide gel electrophoresis or molecular sieve chromatography.

Microorganisms containing the recombinant DNA molecules of this invention are exemplified by cultures deposited in the American Type Culture Collection and identified as follows. Three cultures were deposited on September 10, 1984:

E.coli K12, strain JM83 (pSA304)

E.coli K12, strain JM83 (pSA307)

S.lividans (pSA3721).

Two cultures were deposited on September 27, 1974:

E.coli HB101 (pSAT9724)

S.lividans (pSAT7026).

These cultures have been assigned Accession Numbers \_\_\_\_\_ through \_\_\_\_\_.

While we have hereinbefore presented a number of embodiments of this invention, it is apparent that our basic construction can be altered to provide other embodiments which utilize the processes and compositions of this invention. Therefore, it will be appreciated that the scope of this invention is to be defined by the claims appended

hereto rather than by the specific embodiments which have been presented hereinbefore by way of example.

-36-

We Claim:

1. A DNA sequence characterized in that at least a portion thereof codes for a streptavidin-like polypeptide selected from the group consisting of:
  - (a) SA304, SA307, SA324;
  - (b) DNA sequences encoding polypeptides which hybridize to any of the foregoing DNA sequences and which code on expression for a streptavidin-like polypeptide; and
  - (c) DNA sequences which code on expression for a polypeptide coded for on expression of any of the foregoing DNA sequences.
2. A DNA sequence according to claim 1, characterized in that it contains a sufficient portion of a signal DNA sequence to cause secretion of the streptavidin-like polypeptide.
3. A DNA sequence according to claim 2, characterized in that it contains a sufficient portion of a signal DNA sequence to cause maturation of the streptavidin-like polypeptide.
4. A recombinant DNA molecule comprising a DNA sequence according to claims 1, 2 or 3.
5. A recombinant DNA molecule according to claim 4, wherein said DNA sequence is operatively linked to an expression control sequence in said molecule.
6. A recombinant DNA molecule according to claim 5 wherein the expression control sequence is selected from the group consisting of the E.coli lac system, the E.coli trp system, the E.coli  $\beta$ -lac system, the TAC system, the TRC system, the major operator and promoter regions of bacteriophage  $\lambda$ , the control region of filamentaceous single-stranded DNA phages, the control regions of Streptomyces or other gram positive bacteria and other sequences which control the expression of genes of prokaryotic

-37-

or eukaryotic cells and their viruses or combinations thereof.

7. A recombinant DNA molecule according to claim 6, selected from the group consisting of  
5 pSA304, pSA307 and pSA3721.

8. A host transformed with at least one recombinant DNA molecule according to claim 5, the expression control sequence in that recombinant DNA molecule being operative when linked to a DNA  
10 sequence in said host.

9. A transformed host according to claim 8 wherein the host transformed is selected from the group consisting of strains of Streptomyces lividans, other Streptomyces, other gram positive  
15 bacteria, E.coli, Pseudomonas, Bacilli, yeast, other fungi, animal and plant hosts and human tissue cells.

10. A transformed host according to claim 9, selected from the group consisting of S. lividans(pSA3721), E.coli K12(pSA304) and E.coli  
20 K12(pSA307).

11. A streptavidin-like polypeptide which is produced by a transformed host according to claim 8.

12. A polypeptide characterized in that  
25 it is coded for on expression of a DNA sequence according to claim 1, 2 or 3.

13. A method for producing a streptavidin-like polypeptide comprising the step of culturing a host transformed with a recombinant DNA molecule  
30 according to claim 4.

14. The method according to claim 13, characterized in that the host transformed is selected from the group consisting of strains of Streptomyces lividans, other Streptomyces, other  
35 gram positive bacteria, E.coli, Pseudomonas, Bacilli, yeast, other fungi, animal and plant hosts, and human tissue cells.

15. A hybrid DNA sequence coding for a fused protein, comprising at least two DNA sequences joined end to end and in the same reading frame, at least a portion of said first DNA sequence coding  
5 for a streptavidin-like polypeptide selected from the group consisting of:

(a) SA304, SA307, SA324;

(b) DNA sequences encoding polypeptides which hybridize to any of the foregoing DNA  
10 sequences and which code on expression for a streptavidin-like polypeptide; and

(c) DNA sequences which code on expression for a polypeptide coded for on expression of  
15 said second DNA sequence coding for another protein, polypeptide, peptide or amino acid.

16. A hybrid DNA sequence according to claim 15, characterized in that it contains a sufficient portion of a signal DNA sequence to cause  
20 secretion of the fused protein.

17. A hybrid DNA sequence according to claim 16, characterized in that it contains a sufficient portion of a signal DNA sequence to cause maturation of the fused protein.

25 18. A hybrid DNA sequence according to claim 15, 16 or 17, in which said second DNA sequence codes for tissue plasminogen activator, selected from the group consisting of:

(a) SAT9724, SAT7021;

30 (b) DNA sequences encoding polypeptides which hybridize to any of the foregoing DNA sequences and which code on expression for the fused protein; and

(c) DNA sequences which code on expression  
35 of any of the foregoing DNA sequences.

19. A hybrid DNA sequence according to claim 15, 16 or 17, wherein the second DNA sequence

encodes polypeptides selected from the group consisting of human and animal interferons, human and animal growth hormones, antigens of FMDV, antigens of HBV, human insulin, human blood factors and erythropoietin.

20. A recombinant DNA molecule comprising a hybrid DNA sequence according to claim 15, 16 or 17, wherein said hybrid DNA sequence is operatively linked to an expression control sequence in said molecule.

21. A recombinant DNA molecule according to claim 20, wherein the expression control sequence is selected from the group consisting of the E.coli lac system, the E.coli trp system, the E.coli  $\beta$ -lac system, the TAC system, the TRC system, the major operator and promoter regions of bacteriophage  $\lambda$ , the control region of filamentaceous single-stranded DNA phages, the control regions of Streptomyces or other gram positive bacteria and other sequences which control the expression of genes of prokaryotic or eukaryotic cells and their viruses or combinations thereof.

22. A recombinant DNA molecule according to claim 20, wherein said hybrid DNA sequence contains a second DNA sequence encoding polypeptides selected from the group consisting of human and animal interferons, human and animal growth hormones, antigens of FMDV, antigens of HBV, human insulin, human blood factors, tissue plasminogen activator and erythropoietin.

23. A recombinant DNA molecule according to claim 21, selected from the group consisting of pSAT9724 and pSAT7026.

24. A host transformed with at least one recombinant DNA molecule according to claim 20, the expression control sequence in that DNA molecule

-40-

being operative when linked to a DNA sequence in said host.

25. A transformed host according to claim 24, wherein the host transformed is selected from the group consisting of strains of Streptomyces lividans, other Streptomyces, other gram positive bacteria, E.coli, Pseudomonas, Bacilli, yeast, other fungi, animal and plant hosts and human tissue cells.

26. A transformed host according to claim 24, selected from the group consisting of E.coli HB101 IQ (pSAT9721) and S.lividans (pSAT7026).

27. A fused protein which is produced by a transformed host according to claim 24.

28. A fused protein characterized in that it is coded for on expression of a hybrid DNA sequence according to claim 15, 16 or 17.

29. A fused protein, according to claim 28, wherein the protein, polypeptide, peptide or amino acid encoded by the second DNA sequence is selected from the group consisting of human and animal interferons, human and animal growth hormones, antigens of FMDV, antigens of HBV, human insulin, human blood factors, tissue plasminogen activator and erythropoietin.

30. A fused protein characterized in that it is coded for on expression of a hybrid DNA sequence according to claim 15, 16 or 17.

31. A fused protein which is produced by a transformed host according to claim 24.

32. A method for producing a fused protein comprising the step of culturing a host transformed with a recombinant DNA molecule of claim 20.

33. The method of claim 32, wherein the hybrid DNA sequence contains a second DNA sequence encoding polypeptides selected from the group consisting of human and animal interferons, human and animal growth hormones, antigens of FMDV, antigens

-41-  
of EBV, human insulin, human blood factors, tissue  
plasminogen activator and erythropoietin.

1/10

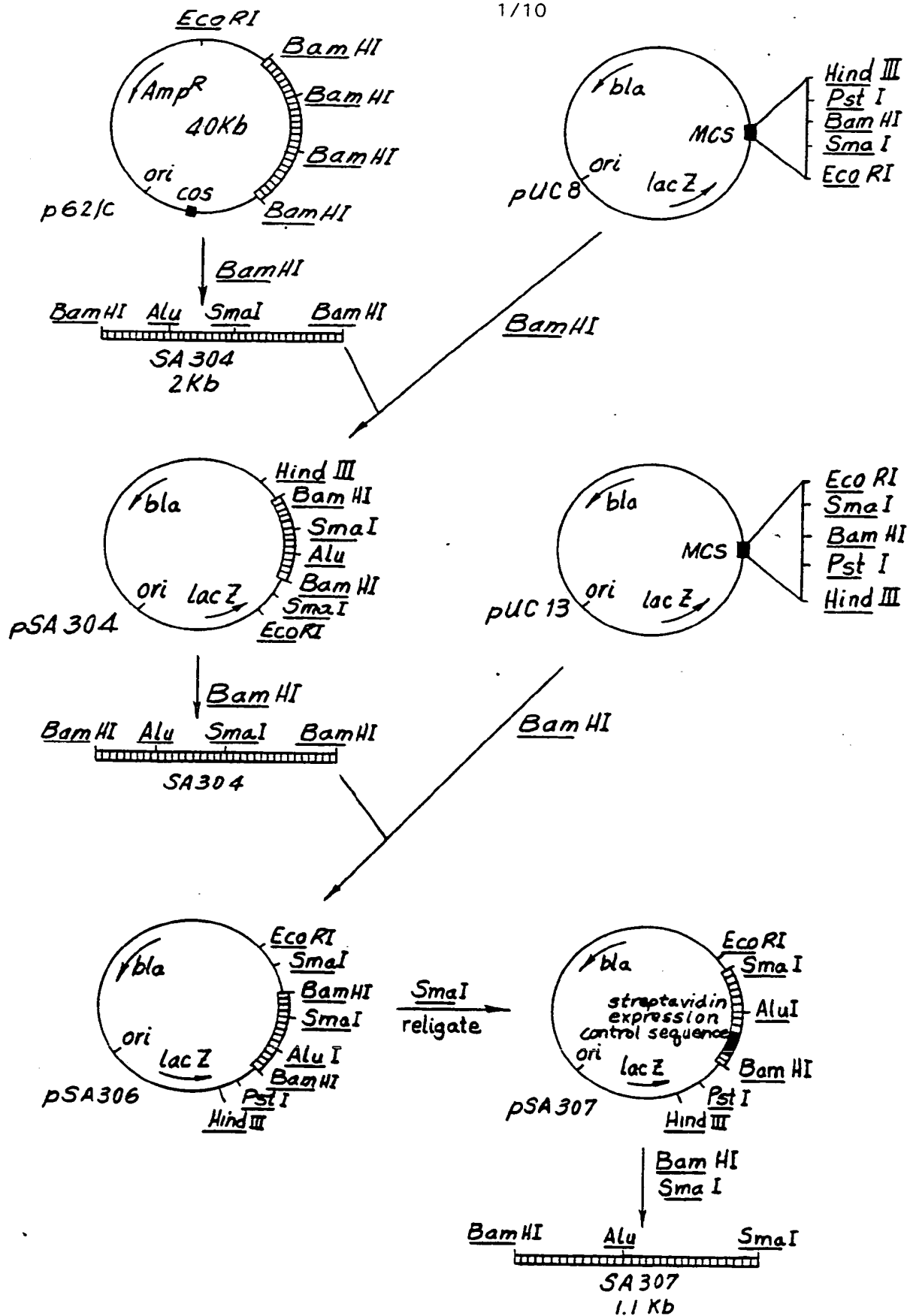


FIG. 1



3/10

FIGURE 2, PAGE 2

301 CCTCGACCATCCAGTTCTGCGGCCAAAGACACATGCGGCACCTGCTGTTTGTTCACCGACACCGTCAGGTGCACGCGCGAGGTCACAAACCTTGACGGGCG  
 GGAGCTGGTAGGTCAAGACGGCGGTTTCTGTGTACGGCGTGACGACAAACAAGTGGCTGTGGCAGTCCACGTGCGCGCTCCAGTGTGGAACCTGCCCCGC 400

401 GGATACGGACGGGCACGCGCCCTCCGTCCCGCGCGGCAACAACCTAGGGGAGTATTTTCGTGCTCAGGTGCAGGATCGTCTGCA  
 CCTATGCCCTGCCGCGTGGGTGTCGCGGGAGGACAGGGGCGCGCCGTTGTTGATCCCTCATAAAAGCACAGAGTGTACGCGTCTAGCAGCAACGT 500

501 GCCATCGCGGTTTCCCTGACACAGGTCTCGATTACGGCCAGCGCTTCGGCAGACCCCTCCAGGACCTCGAAGGCCCGAGGCTCTCGCGCGCGAGGCGGCA  
 CCGTAGCGGCAAGGGACCTGGTCCAGAGCTAATGCGGTGCGGAAGCGCTCTGGGAGGTTCTTGAGCTTCGGGTTCCAGAGCGCGGCTCCGGCCGT 600

A P S P F P P P R S R L R P A L R Q T P P R T R R P R S R P P R P A  
 B H R R F P D H G L D Y G Q R F G R P L Q G L E G P G L G R R G R H  
 C A I A V S L T T V S I T A S A S A D P S K D S K A Q V S A A E A G I

BH SGXH  
 PIMA  
 LAEE  
 2133

M N L I  
 T T H I  
 H P H I  
 F NN LU A4 3H  
 F NN LU A4 3H

H H H H  
 SINHHI  
 SNUHHN  
 HPDAAP  
 212111

H H H H  
 I H N H P A  
 1 1 1 1

S HMNC  
 PSCR  
 APIF  
 2111

M N L I  
 H IMHN M D  
 NSHL B P  
 PTAA O N  
 1113 1 1

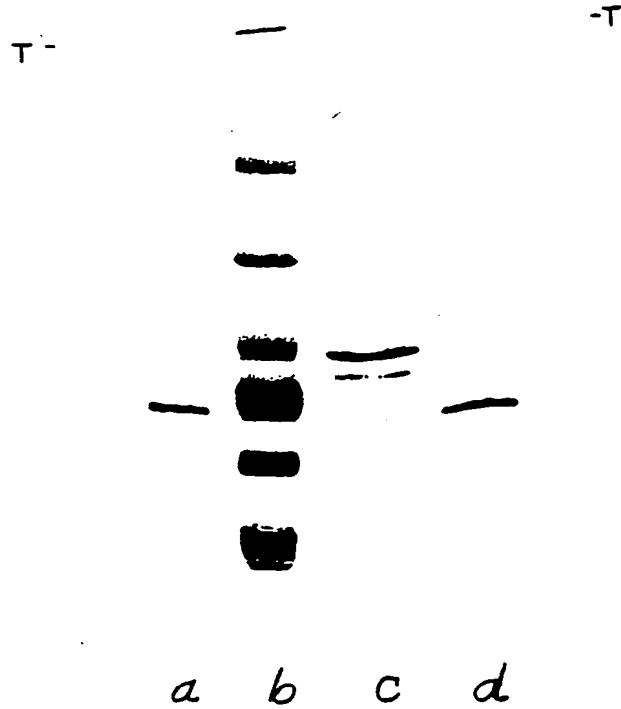
C A R S S L Q  
 A Q D R R C S  
 M R K I V V A

F N U 4 H  
 BSHHMN  
 GFPAPSA  
 LIHEAPE  
 1113211





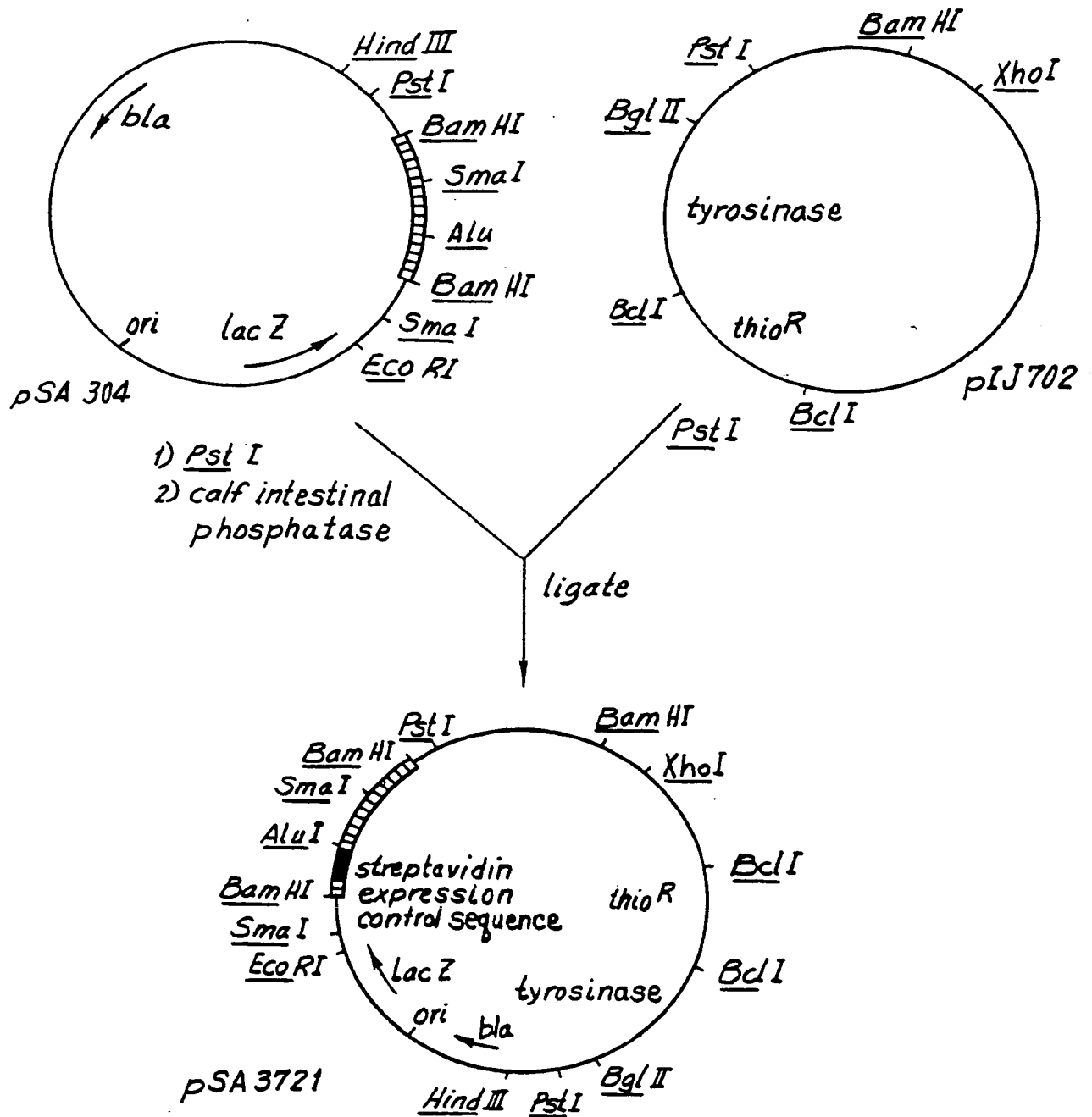
6/10



**FIG.3**

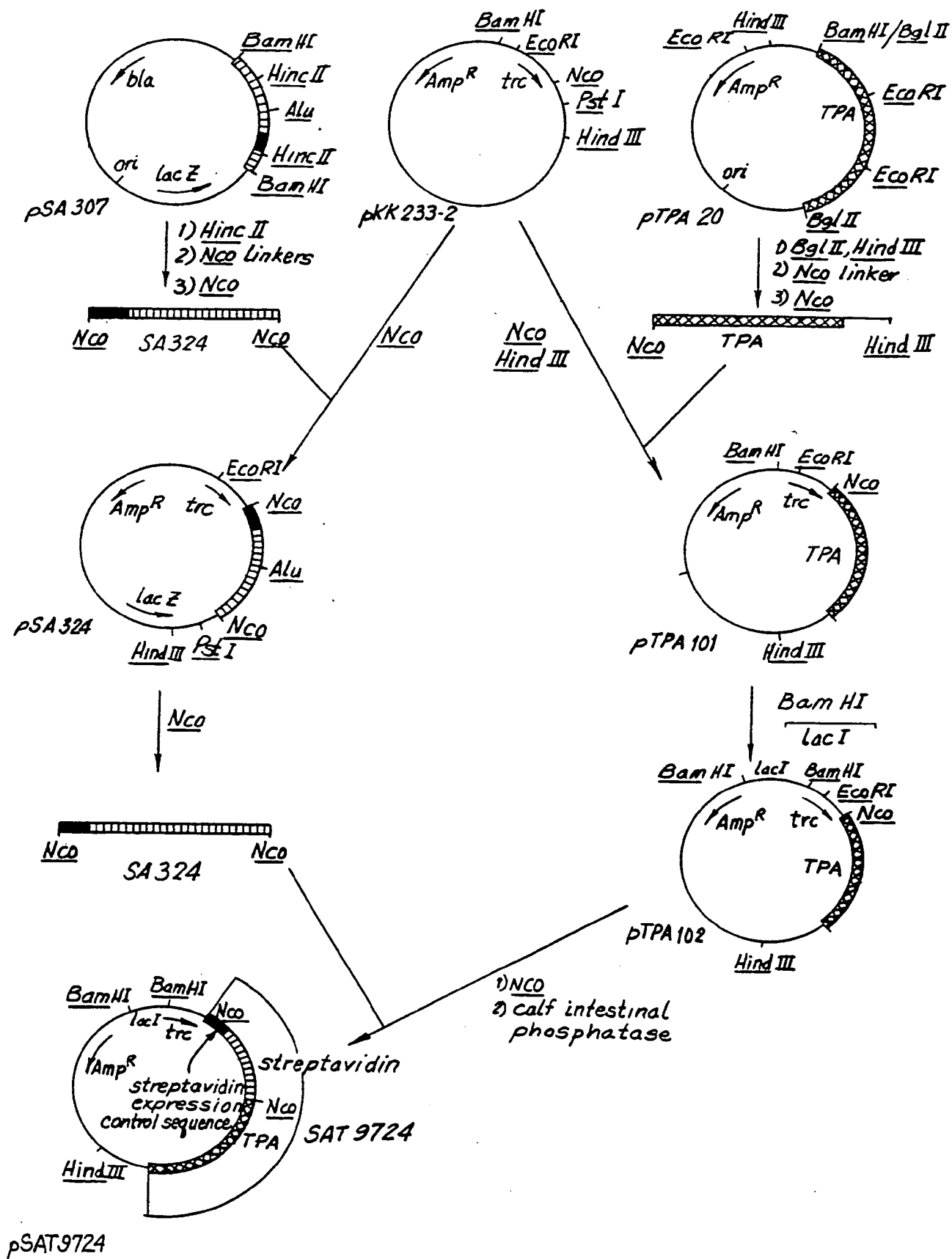
7/10

FIG. 4



8/10

FIG. 5



9/10

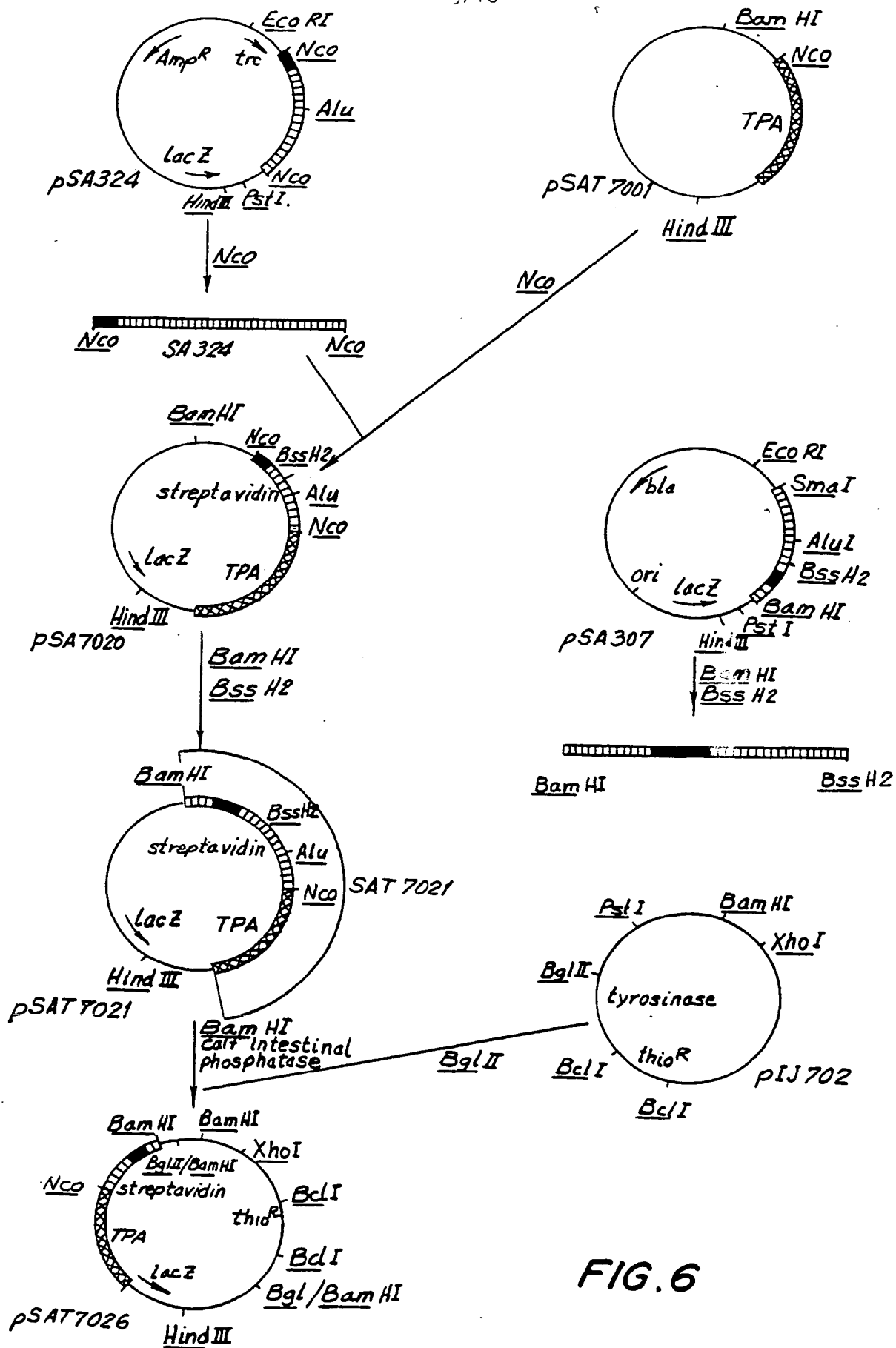


FIG. 6

10/10

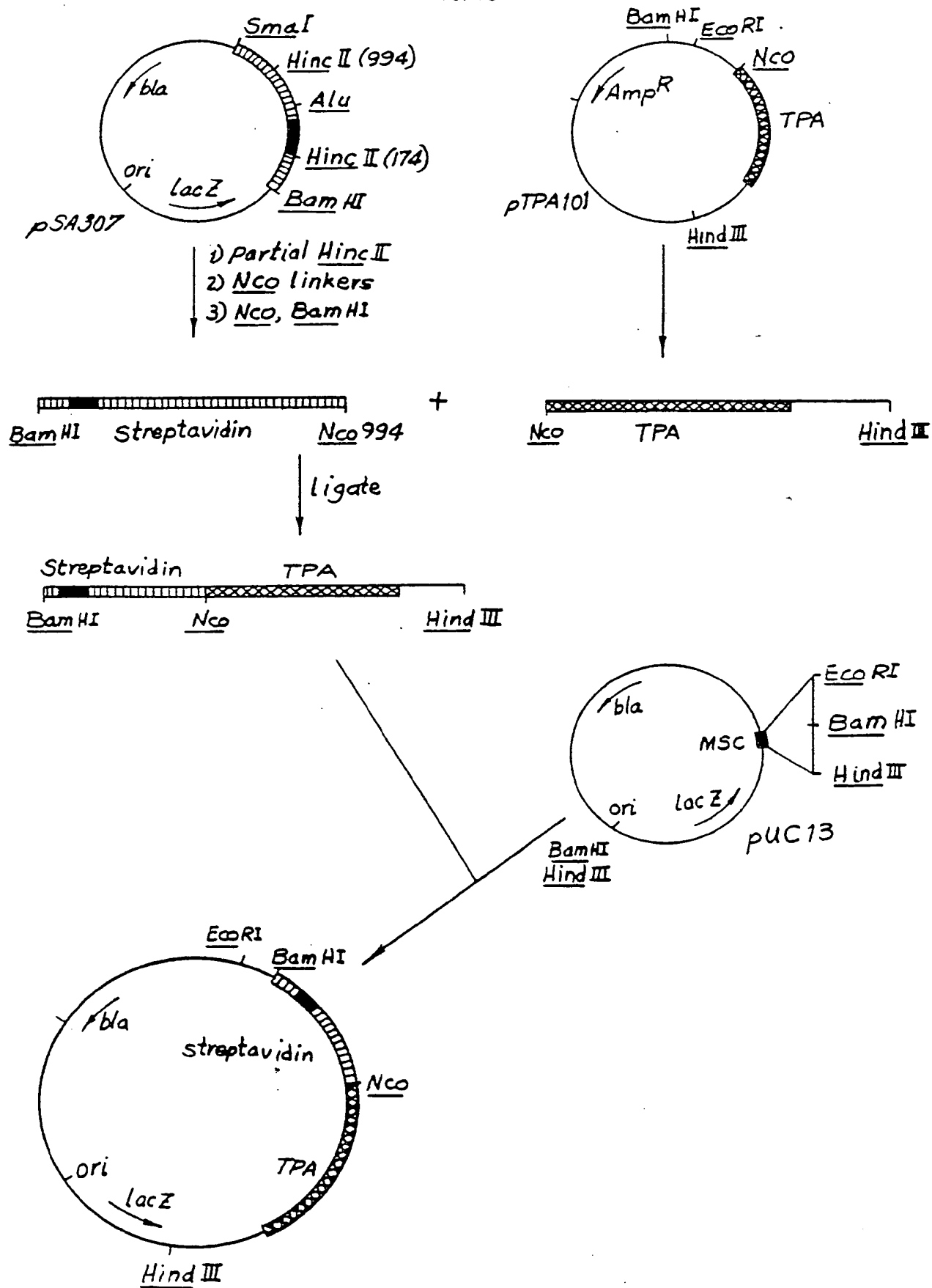


FIG. 7

# INTERNATIONAL SEARCH REPORT

International Application No PCT/US85/01901

<b>I. CLASSIFICATION OF SUBJECT MATTER</b> (If several classification symbols apply, indicate all) <sup>1</sup>		
According to International Patent Classification (IPC) or to both National Classification and IPC I.P.C. : C07H 21/04; C12N 1/00; C07K 15/00; C12P 21/00; C12P 21/02		
<b>II. FIELDS SEARCHED</b>		
Minimum Documentation Searched <sup>4</sup>		
Classification System	Classification Symbols	
U.S.	435/68,70,71,91,172.3,243,253,255,256,317; 47/1.4; 260/112.5R, 112R; 514/2; 536/27; 935/10,11,22,29,33,38,39,47,48,66-75	
Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched <sup>5</sup>		
CA SEARCH DATABASE: 1967-85		
<b>III. DOCUMENTS CONSIDERED TO BE RELEVANT</b> <sup>14</sup>		
Category <sup>6</sup>	Citation of Document, <sup>16</sup> with indication, where appropriate, of the relevant passages <sup>17</sup>	Relevant to Claim No. <sup>18</sup>
Y	N, Rosenberg et al, Ann. Rev. Genet., Vol. 13, 1979, pages 319-353.	1-33
Y	N, <u>Molecular Biology of The Cell</u> Alberts et al, Garland Publishing, Inc., 1983, New York, pages 177-178.	1-33
Y	N, Edlund et al, Proc. Natl. Acad. Sci. USA, Vol. 80, 1983 pages 349-352	1-33
Y	N, DeBoer et al, <u>Promoter, Structure and Function</u> , Rodriguez et al (ed.), Praeger Publishers, 1982, New York, pages 462-481	1-33
<div style="display: flex; justify-content: space-between;"> <div style="width: 48%;"> <p><sup>15</sup> Special categories of cited documents:</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> </div> <div style="width: 48%;"> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.</p> <p>"&amp;" document member of the same patent family</p> </div> </div>		
<b>IV. CERTIFICATION</b>		
Date of the Actual Completion of the International Search <sup>2</sup>		Date of Mailing of this International Search Report <sup>3</sup>
09 December 1985		27 DEC 1985
International Searching Authority <sup>1</sup>		Signature of Authorized Officer <sup>10</sup>
ISA/US		James Martinelli

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)		
Category *	Citation of Document, <sup>16</sup> with indication, where appropriate, of the relevant passages <sup>17</sup>	Relevant to Claim No <sup>18</sup>
Y	N, Taniguchi et al, Nature, Vol. 285 1980, pages 547-549	1-33
Y	N, Roberts et al, Proc. Natl. Acad. Sci. USA, Vol. 76, 1979, pages 760-764	1-33
Y	N, Gray et al, Nature, Vol. 295, 1982 pages 503-508	1-33
Y	N, Burrell et al, Nature, Vol. 279, 1979, pages 43-47	1-33
Y	US, A, 4,338,397, Published 6 July 1982, Gilbert et al	1-33
Y	US, A, 4,349,629, Published 14 September 1982, Carey et al	1-33
Y	N, Nagata et al, Nature, Vol. 284, 1980, pages 316-320	1-33
Y	N, Kupper et al, Nature, Vol. 289, 1981, pages 555-559	1-33
Y	N, Wallace et al, Nucleic Acids Res., Vol. 9, 1981, pages 879-894	1-33

## FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

Y	US, A, 4,332,892, Published 1 June 1982, Ptashne et al.	1-33
Y	N, Seeburg, DNA, Vol. 1, 1982, pages 239-249	1-33
Y	N, Hofmann et al, Proc. Natl. Acad. Sci. USA, Vol. 77, 1980, pages 4666-4668	1-33

V. ☐ OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE <sup>10</sup>

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. ☐ Claim numbers ..... because they relate to subject matter <sup>12</sup> not required to be searched by this Authority, namely:

2. ☐ Claim numbers ..... because they relate to parts of the International application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out <sup>13</sup>, specifically:

VI. ☒ OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING <sup>11</sup>

This International Searching Authority found multiple inventions in this International application as follows:

- I. Claims 1-10, 13-26, 32 and 35 classified in 536/27 and 435/68, 243, 253, 255 and 256 and
- II. Claims 11, 12 and 27-31 classified in 260/112R and 260/112.5R.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the International application for which fees were paid, specifically claims:

3. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4. ☐ As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

## Remark on Protest

- ☐ The additional search fees were accompanied by applicant's protest.
- ☒ No protest accompanied the payment of additional search fees.